

Appl. No. 09/823,837
Amtd. dated June 22, 2004
Reply to Office action of March 30, 2004

Amendments to the Claims:

This listing of claims will replace all prior versions, and listings, of claims in the application:

Listing of Claims:

1. (Original) A method of scheduling in a mixed workload environment on a computing system having a CPU resource and a permanent storage resource, the computing system servicing requests from one or more clients, comprising:

executing a current process on the CPU resource and the storage resource, the current process having been dispatched to service a current client request;

performing a contention check while executing the current process to determine whether a new client request has a transaction priority that is greater than the transaction priority of the current client request;

if the transaction priority of the new client request is greater than that of the current request, dispatching a process to service the new client request;

if the transaction priority of the new request is not greater than that of the current request, determining whether the transaction priority of the current request is less than a predetermined threshold priority;

if the transaction priority of the current client request is lower than the predetermined threshold priority and there is higher priority I/O activity present on the storage resource:

delaying the servicing of the current client request and forgoing the servicing of any read aheads for the current client request; and

dispatching a process to service the highest priority client request that is available for service; and

if the transaction priority of the current client request is greater than the predetermined threshold or the priority of the current client request is lower than the predetermined threshold and there is no higher priority I/O activity present on the storage resource:

Appl. No. 09/823,837
Amdt. dated June 22, 2004
Reply to Office action of March 30, 2004

determining whether the current client request requires any read
aheads;

dispatching one or more helper processes to service any required
read aheads; and

returning to the current process to service the current client request.

2. (Original) A method of scheduling in a mixed workload environment as
recited in claim 1,

wherein a maximum priority in the system is 255 and a minimum priority is
1; and

wherein the threshold priority is 151.

3. (Original) A method of scheduling in a mixed workload environment as
recited in claim 1, wherein the step of delaying the servicing of the current client
request includes

delaying the servicing of the current client request by an amount of time
that depends on the transaction priority of the current client request, higher priority
requests being delayed less than lower priority requests, and the amount of the
delay being the sum of a fixed delay and a priority dependent delay.

4. (Original) A method of scheduling in a mixed workload environment as
recited in claim 3,

wherein a maximum transaction priority in the system 255 and a minimum
priority is 1 and the threshold priority is 151;

wherein the fixed delay is about 0.2 seconds; and

wherein the priority dependent delay is the product of a constant and the
difference between the threshold priority and the priority of the current client
request.

5. (Original) A method of scheduling in a mixed workload environment as
recited in claim 4, wherein the constant is approximately 0.02.

Appl. No. 09/823,837
Amdt. dated June 22, 2004
Reply to Office action of March 30, 2004

6. (Original) A method of scheduling in a mixed workload environment as recited in claim 1, wherein the step of delaying the servicing of the current client request includes delaying the servicing by a fixed delay.

7. (Original) A method of scheduling in a mixed workload environment as recited in claim 6, wherein the fixed delay is approximately 10 milliseconds.

8. (Original) A method of scheduling in a mixed workload environment as recited in claim 1, wherein the step of performing a contention check occurs once every time a physical block is transferred from the storage resource.

9. (New) A system, comprising:
a central processing unit ("CPU");
an interface coupled to the CPU and adapted to receive client requests,
said CPU further adapted to:

execute a current process on the CPU, the current process dispatched to service a current client request;

determine whether a new client request has a transaction priority that is greater than the transaction priority of the current client request;

if the transaction priority of the new client request is greater than that of the current request, to dispatch a process to service the new client request;

if the transaction priority of the new request is not greater than that of the current request, to determine whether the transaction priority of the current request is less than a threshold;

if the transaction priority of the current client request is lower than the threshold and higher priority input/output ("I/O") activity is present on a storage resource, to delay the servicing of the current client request and to dispatch a process to service the highest priority client request that is available for service;
and

if the transaction priority of the current client request is greater than the threshold or the priority of the current client request is lower than the threshold

Appl. No. 09/823,837
Amdt. dated June 22, 2004
Reply to Office action of March 30, 2004

and no higher priority I/O activity is present, to determine whether the current client request requires any read aheads and to dispatch one or more helper processes to service any required read aheads.

10. (New) The system of claim 9 wherein the CPU is further configured to delay the servicing of the current request by a variable amount.

11. (New) A system, comprising:

a central processing unit ("CPU") adapted to receive client requests and to:
execute a current process dispatched to service a current client request;

determine whether a new client request has a transaction priority that is greater than the transaction priority of the current client request;

if the transaction priority of the new request is not greater than that of the current request, to determine whether the transaction priority of the current request is less than a threshold; and

if the transaction priority of the current client request is lower than the threshold and higher priority input/output ("I/O") activity is present on a storage resource, to delay the servicing of the current client request and to dispatch a process to service the highest priority client request that is available for service.

12. (New) The system of claim 11 wherein if the transaction priority of the new client request is greater than that of the current request, the CPU is adapted to dispatch a process to service the new client request.

13. (New) The system of claim 11 wherein if the transaction priority of the current client request is greater than the threshold or the priority of the current client request is lower than the threshold and no higher priority I/O activity is present, the CPU is adapted to determine whether the current client request requires any read aheads and to dispatch one or more helper processes to service any required read aheads.